Biophysical Chemistry 199 (2015) 39–45

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

A novel data-driven algorithm to reveal and track the ribosome heterogeneity in single molecule studies

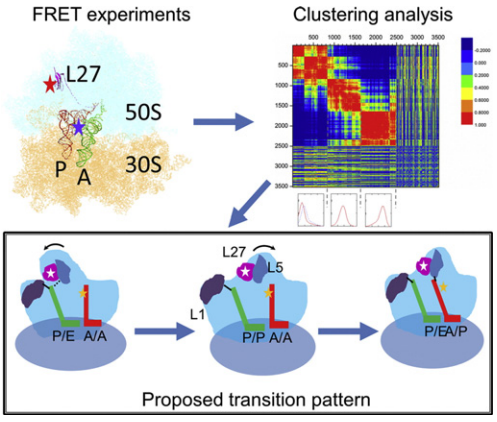
Haopeng Yang, Ming Xiao [1], Yuhong Wang *

University of Houston, 4800 Calhoun Rd, Houston, TX 77214, USA

HIGHLIGHTS

- The FRET signals at the single ribosome's peptidyl transfer center are studied.
- Inhomogeneous subpopulations are identified with a novel data-driven algorithm.
- The subpopulation exchange among each other spontaneously in 2 min time intervals.
- The exchange patterns suggest that the ribosomal and tRNA dynamics are correlated.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form 20 February 2015
Accepted 23 February 2015
Available online 3 March 2015

Keywords:
Ribosome conformational dynamics
Single molecule FRET
Data-driven algorithm
Dynamic heterogeneity

ABSTRACT

The unique advantage of the single molecule approach is to reveal the inhomogeneous subpopulations in an ensemble. For example, smFRET (single molecule fluorescence resonance energy transfer) can identify multiple subpopulations based on the FRET efficiency histograms. However, identifying multiple FRET states with overlapping average values remains challenging. Here, we report a new concept and method to analyze the single molecule FRET data of a ribosome system. The main results are as follows: 1. based on a hierarchic concept, multiple ribosome subpopulations are identified. 2. The subpopulations are self-identified via the cross-correlation analysis of the FRET histogram profiles. The dynamic heterogeneity is tracked after 2 min intervals on the same ribosomes individually. 3. The major ribosome subpopulations exchange with each other with a certain pattern, indicating some correlations among the motions of the tRNAs and the ribosomal components. Experiments under the conditions of 20% glycerol or 1 mM viomycin supported this conclusion.

1. Introduction

Single molecule methods track individual molecules in single reaction trajectories. Therefore, inhomogeneous subpopulations and rare reaction pathways can be observed in complex and out of phased biological systems, such as the ribosome system [1]. However, due to intrinsic fluctuations and instrumental noises, it is not always straightforward to

* Corresponding author. Tel.: +1 713 743 7941; fax: +1 713 743 8351.
E-mail address: ywang60@uh.edu (Y. Wang).
[1] Present addresses: Science China Press, Beijing 100717, China.

http://dx.doi.org/10.1016/j.bpc.2015.02.008
0301-4622/

extract kinetics and conformational information from a single molecule data. For example, in the single molecule fluorescence resonance energy transfer (smFRET), the subpopulations can be sorted by the FRET efficiency histograms or via the hidden Markov analysis for kinetic information [2]. Both methods are model-dependent and adding more states always helps them fit the data better. When the distributions of multiple FRET states are overlapped, the fitting becomes more biased by the initial parameters. Consequently, the kinetic information will not be reliable because it depends on the identification of FRET states. Here, we report a new model-free sorting method to identify the ribosome subpopulations.

The ribosome moves processively on the mRNA to assemble polypeptide. The aminoacyl tRNA pairs with the mRNA codon at the A-site and accommodates into the ribosome. Then, the elongating peptide chain on the P-site tRNA is transferred to the A-site tRNA via the new amide bond between the peptide chain and the aminoacyl tRNA. Afterward, the tRNAs move from the A- and P-sites to the P- and E-sites, respectively, which moves the next codon into the A-site for a new elongation cycle. The cycles continue until a stop codon moves into the A-site to finish the nascent peptide synthesis. Multiple ribosomal components and the tRNAs undergo large conformational changes during these processes [3,4]. The two subunits (30S and 50S) ratchet relative to each other; the 30S head swivels, and the tRNAs fluctuate between classical and hybrid binding positions in the 50S [5–9]. A cryo-EM study has identified more than 30 different subpopulations in the ribosome pre-translocation complex with different combinations of these movements, implying that the ribosome allosteric motions are not strongly correlated. On the other hand, it appears that more extensive inter-subunit ratcheting enables the tRNAs sample more conformations near their binding pockets, suggesting that the ribosome scaffolds dictate the tRNA dynamics [10]. In the meantime, we reported single molecule FRET studies on the ribosome that probed the distance between the A- or P-site tRNAs and the ribosomal protein L27. Three FRET states were observed at values of approximately 0.2, 0.47 and 0.68. However, the ribosomes cannot be simply assigned to only three states because individual ribosomes sample the three FRET states differently. Considering both values and dynamics of the FRET states, at least seven ribosome subpopulations can be identified: three subpopulations sample only one of the FRET states; three subpopulations sample only two of the three FRET states; and one subpopulation samples all three FRET states [11,12]. Therefore, more heterogeneity can be deduced, although not directly observable. This analysis method sorts the ribosome subpopulation based on a series of FRET data points instead of just one data point, therefore, the subpopulations identified are at a higher level of heterogeneity than the FRET values, which we have named the hierarchic model. The main difference between this hierarchic model and other methods is that both FRET values and dynamics are considered to define a subpopulation, whereas most other research reports define subpopulations only by the FRET values [13]. The hierarchic model agrees with the above-mentioned cryo-EM study that shows that the ribosome inter-subunit ratcheting conformations affect the tRNA dynamics. Similarly, the subpopulations identified by the hierarchic model are ribosomes with different conformations that allow different tRNA dynamics. However, our previous analysis still defines the FRET states via Gaussian fittings to the FRET efficiency histograms, which are biased by the initial guess values. In this report, the subpopulations are self-identified via cross-correlation analysis of the FRET efficiency histogram profiles. This analysis, therefore, eliminated the model-fitting of FRET states.

2. Results

2.1. Single molecule FRET experimental conditions

The FRET-paired dyes (Cy3–Cy5) are tethered on the C53 residue of the ribosomal protein L27 and the D-loop of the tRNA (D16/17), respectively, as described previously [11,14] (Fig. S1 and sample preparation description in the Supplemental material). In the pre-translocation complex, the dye-labeled $tRNA^{Phe}$ carrying the fMet–Phe dipeptidyl chain is at the A-site, while the vacant $tRNA^{fMet}$ is at the P-site (Fig. S1). The complex is tethered to the surface via the surface–streptavidin–biotin–mRNA interaction. A 532 nm laser is directed through a TIRF objective (Nikon Instruments) to illuminate the sample above the total internal reflection angle and generates an evanescent wave to excite the donor dye. The fluorescence emitted from the dyes are collected by the same objective and imaged on the CCD camera (Cascade II, Photometrics). The FRET efficiencies are calculated according to $E = I_{acceptor} / (I_{donor} + I_{acceptor})$. To remove signals with photophysical flickering, the fluctuation of the total fluorescent intensity ($I_{donor} + I_{acceptor}$) has to obey the Poisson distribution to warrant no donor photophysical flickering [15]. The acceptor physical flickering is tolerated because it occurs at a much lower frequency than the observed inhomogeneity [15]. Meanwhile, the signal noise level is set to less than 20%. For every ribosome, the fluorescence intensities are collected for 3 segments at 2 min intervals. Every segment contains 200 data points at 100 ms time resolution (Fig. 1, top panel). The FRET efficiencies are calculated (Fig. 1, second panel) and binned in 20 intervals between 0.02 and 1.02 to generate the histogram (Fig. 1, third panel). In the 2 min intervals, we have observed subpopulation transitions, which means that these conformational changes occur at the order of minutes. This time scale is slow compared with the 10–20 s^{-1} peptide elongation with elongation factors, but is fast compared with a factor-free spontaneous protein synthesis ($<0.03\ min^{-1}$) [16,17].

2.2. Identify the subpopulations

The FRET efficiency histograms obtained from individual ribosomes exhibit large heterogeneity. Fig. 2 displays 6 ribosome traces of the first time segment. The traces in the top panel prefer the low FRET state and sample the other FRET states much less frequently; the traces in the bottom panel sample the high FRET states most of the time. In a conventional analysis, these traces are pooled together, and an overall Gaussian-fitting of the total FRET efficiency histogram will generate several FRET states (Fig. S2). This analysis method is model-dependent and has ignored the heterogeneity of every trace. On the other hand, although no two traces are exactly the same, it is obvious that traces can be grouped based on the similarity of the FRET histogram profiles, such as the two panels in Fig. 2. This is reasonable because the FRET signal reflects the dynamics of the A-site tRNA that are dependent on the ribosome scaffolds. Therefore, similar FRET histogram profiles imply similar ribosome scaffolds. For example, in the ratcheted and non-ratcheted ribosome, the tRNA prefers the hybrid (high FRET state) and classical (low FRET state) binding sites respectively (Fig. S1 shows the expected FRET values and the tRNA configurations based on X-ray information [11]). Accordingly, ribosomes generating similar FRET histogram profiles are grouped as one subpopulation. Here, the definition of a subpopulation is not only the FRET values, but also the sampling frequencies among the states that indicate the relative energy gaps. The FRET dynamics are measured at 100 ms interval for 20 s (200 data points). Therefore, the subpopulations defined here are at a higher level than the FRET states only, and include the heterogeneity of single traces.

The FRET histogram profile similarity is quantified by the Pearson's cross-correlation coefficients among traces as shown in Fig. 2. Fig. 3 displays plots of the coefficient matrices of ribosomes under three experimental conditions: (a) in aqueous buffer solution (3519 traces); (b) in 20% glycerol buffer solution (1908 traces); (c) in the presence of viomycin in aqueous buffer (673 traces). For example, Fig. 3a is a matrix of 3519×3519 elements. For each element $M_{i,j}$, the value is the Pearson's correlation between the "ith" and "jth" traces. Therefore, the matrix is symmetric along the diagonal line because M_{ij} equals M_{ji}, and the values